Vegetative Assessment in an Urban Environment

By Alan Walters, Agricast

AGRICAST

Vegetative Assessment in an Urban Environment

Final Report

Contract Number 4600000983

Presented to the City of Clovis, CA
and the
Department of Water Resources, San Joaquin District

Alan Walters

December 3, 2001

12/03/01 visual information 1

Agricast is a small business, incorporated in the State of California, located at Escondido, CA. We are a satellite imagery distributor for Space Imaging and for Earthscan. Our principal product is image processing and not necessarily sale of raw, unprocessed imagery. We specialize in work where satellite imagery is used for land-use applications. Our address and contact information are shown below:

P.O. Box 22

Escondido, CA 92033-0022

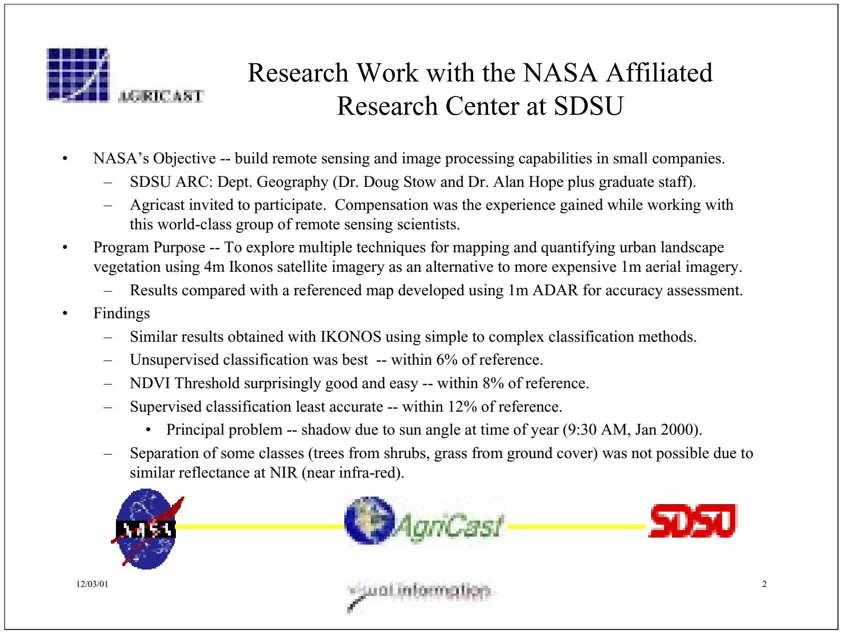
Voice (760) 480-7884

Fax (760) 480-1115

The work described here is a follow-on to research performed at the NASA Affiliated Research Center, San Diego State University, Department of Geography. It was during this research that image processing techniques necessary to measure irrigated vegetation were developed and refined.

This report describes the underlying research, the work performed for the City of Clovis and DWR, along with a description of how the work was performed and the processes used.

1



The IKONOS satellite (owned by Space Imaging) was launched in September 1999 and provides the first commercial source of high-resolution satellite imagery. Previous to IKONOS, work requiring high resolution imagery was accomplished using expensive digital aerial imagery at three and four times the price compared to this new satellite source. Since there were no examples of what could be accomplished with this new source of imagery, NASA sponsored a number of research projects through its Affiliated Research Centers (ARC), such as the Department of Geography at San Diego State University.

Agricast has a growing reputation of quality work in the area of land-use applications using satellite imagery and was invited to participate as the "visiting researcher". Our compensation was the opportunity of refine existing image processing skills and learn new ones working under the supervision of a world class remote sensing faculty.

This project required six months of concentrated activity which included: field work, image processing, use of GIS, review and critique of results at SDSU, exploration of alternative methods at SDSU, and then back to the field for verification - validation.

Results of this program are shown on the slide with the principal finding that if one can be satisfied with an estimate of irrigated vegetation without knowing the proportion of trees and shrubs, to grass and ground cover -- a simple, quick NDVI slice is surprisingly accurate.

Ways to Classify a Scene

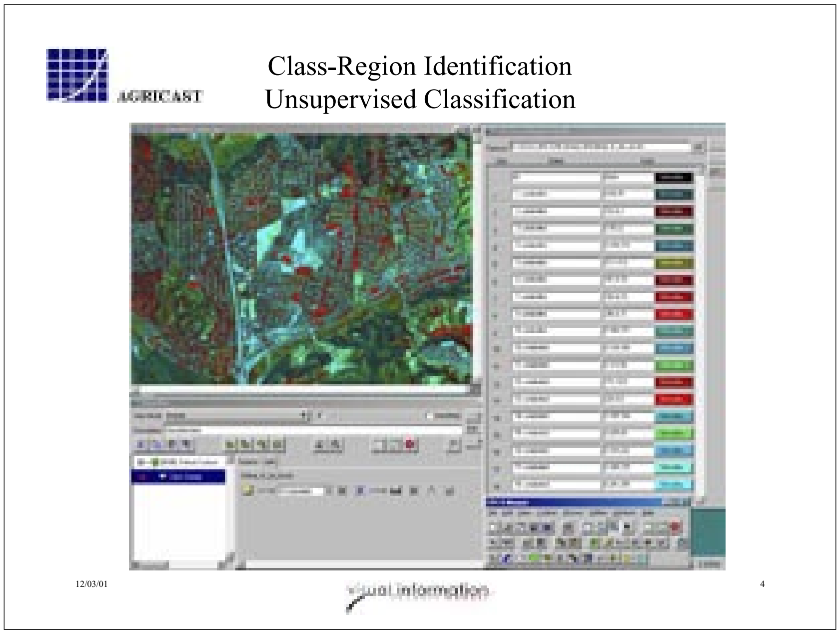
- **Unsupervised classification** -- uses an image processing algorithm in which multiple classes (100 in these cases) are selected for output. These classes are then visually identified, labeled and aggregated into landscaped vegetation or impervious classes. Requires good image interpretation skills. Process is time consuming.
- **Supervised classification** -- uses color coded example (training) areas identified within the scene to represent various feature classes. An image processing algorithm uses the spectral reflectance of each sample area to identify by color code all other similar reflectancies in the scene. Requires good image interpretation skills and meticulous selection of example areas. Can yield rapid results, but iteration often required until results agree with ground truth.
- **Normalized Difference Vegetation Index** (NDVI) is computed from the red and near-infrared (NIR) wavebands according to the following equation: NIR - red / NIR + red.
  – A threshold value is placed in the formula where pixels with values below the threshold are classified as impervious; values above the threshold are classified as landscaped vegetation.
  – The NDVI threshold value for the Del Mar area of 0.14 was determined through iterative visual assessment.
  – Requires good image interpretation skills but iteration process is quick and simple.

12/03/01 3

The classic way to classify features shown in satellite imagery is through either supervised or unsupervised classification. In an unsupervised classification, image processing software is used to separate reflectance in the scene into multiple classes -- 100 in the cases noted here. The next step is to visually identify each class, assigned a name, then consolidated with like classes by color. As an example, a highly reflective roofing material might spread across five or more classes. Each would be assigned the name "roof" and the same color.

Supervised classification uses example areas (training sets) to tell the image processing software what to look for. For example, a colored polygon would be used to identify "trees"; a second polygon of another color would be used to identify grass, and so-on until a training set has been established for each major feature in the scene. To limit the number of features, areas not germane to the analysis are often eliminated for the scene. From here the process is automated within the software, resulting in a color coded map showing each feature.

The Normalized Difference Vegetative Index is simply a comparison of the near infrared reflectance of healthy vegetation (band 4) with soil or other impervious surface (band 3). The result is shown in grayscale or can be color-coded. A threshold value may be entered into the NDVI formula to act as a cut-off, showing only values above the index. Quick, easy, and accurate.



The mechanical process of unsupervised process is shown here using Earth Resources Mapping image processing software. One must manually identify each class. This shows a scene divided into 18 classes. The writer normally commences at the bottom of the class index, turns the color automatically chosen by the software into a bright pink (different from everything else in the scene), and then visually identifies the feature. Once this is accomplished, the feature is named and assigned a color.

Please note that a feature can consist of more than one class due to different types of roofing materials, street paving, grass textures, tree densities, etc. One must keep the objective in mind when performing an unsupervised classification.

Here, the objective is the identification of irrigated vegetation by area. Imagine worn areas in a school yard or park. It would be perfectly accurate to classify these areas as dirt. But, dirt areas might not be included in the roll up of irrigated areas. This class might better be used to describe areas of raw earth. Consequently it would be better to classify these worn areas in play ground and parks as "grass" because these areas are irrigated along with the surrounding healthier areas.

Even so, there is no clear rule-of-thumb. It is the experience and image interpretation skills of the image processor that make the difference.

4

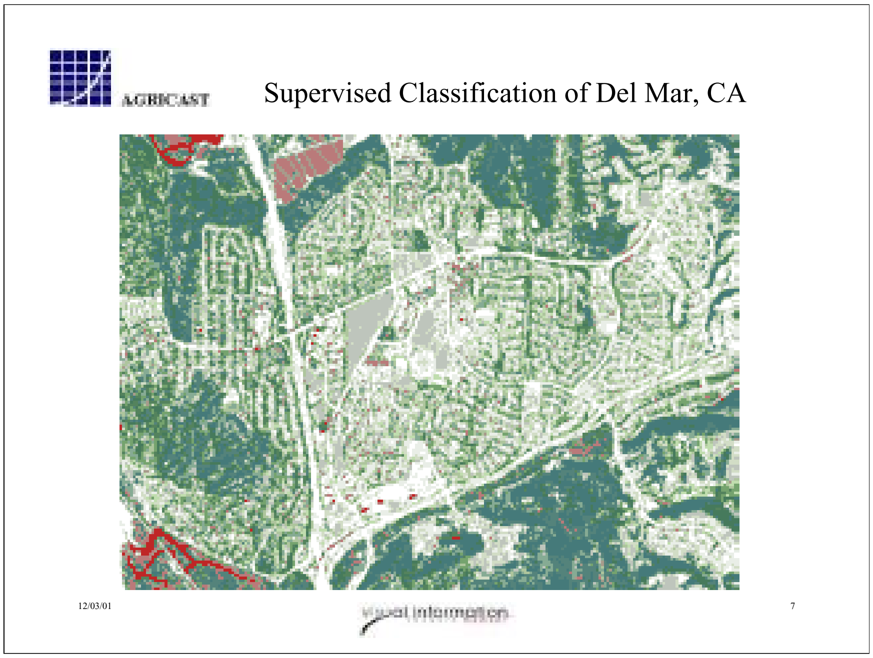
This shows how multiple classes are collapsed to represent the different cover types. At the bottom, the bare dirt areas are color coded gray and identified as dirt. The streets and roof tops are color coded white and identified as impervious. Grass is color coded green, shrubs - light brown, trees dark green and red is used to identify brush.

5

As mentioned before, supervised classification uses example areas (training sets) to tell the image processing software what to look for. Shown here are example areas for impervious (streets and roofs) in white, barren areas in gray, brush in blue-gray, grass and ground cover in green, grass and dirt in yellow, trees and shrubs in dark green, agriculture (at the upper left center) in orange, and shadow in red.

From here the process is automated within the software. First calculate the statistics for the training areas and the run the classification algorithm. This results in the color coded map shown on the next slide.

6



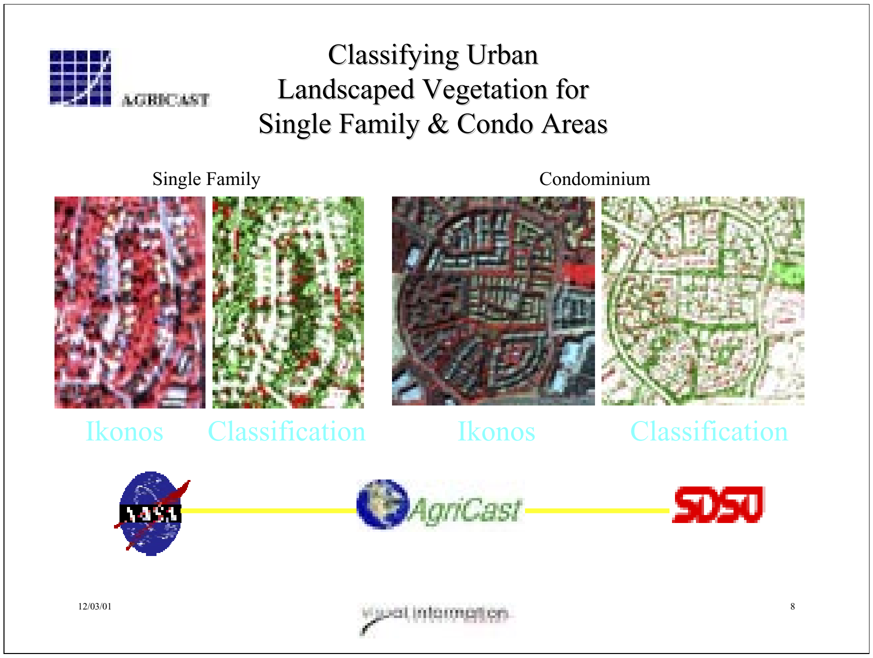
And here is the result of the classification. One can look at the statistics and see an acreage count for the total scene (All), and acreage for each of the classified features in the scene.

However, note that while red was used to depict shadow as can been seen on the northwest sides of buildings at lower left of center, and along the northwest side of trees along streets, we have other red areas which are not shadow.

This is an example of confusion where water in the lagoons has the same dark spectral reflectance as dark areas in shadow. The result, misclassification. Subsequently, all areas not germane to the study were masked -- that is excluded from the scene -- leaving just the residential areas.

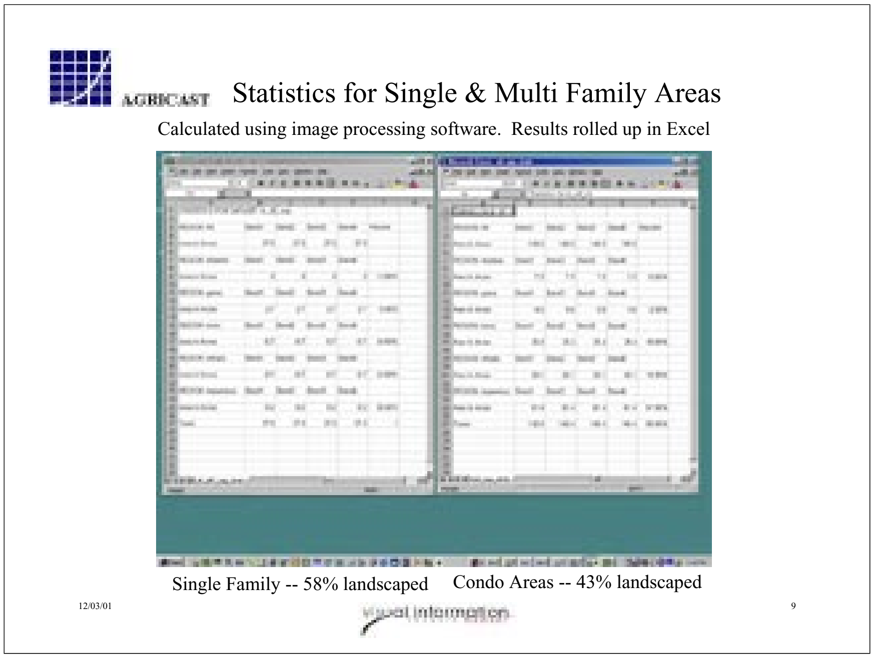
Then, we need not deal with the lagoons, the brush areas, and the barren areas under development. Masking allows focus on just what we were interested in: grass and groundcover, trees and shrubs, dirt and grass (to account for worn spots in play grounds and parks) combinations under irrigation, and other impervious surfaces such as roofs and streets.

7



Two areas were selected for detailed examination; the single family area at the left and an area of dense condominiums at the right. The object was to see if "rules of thumb" could be developed for application to other similar areas.

Additionally these areas were used as ground truth to measure the accuracy of the various classification methods used. This was accomplished by using ADAR multispectral 1m imagery to develop a land cover map. This map was verified and corrected through on the spot ground observation -- pixel by pixel. The result was used as "reference" for evaluation.

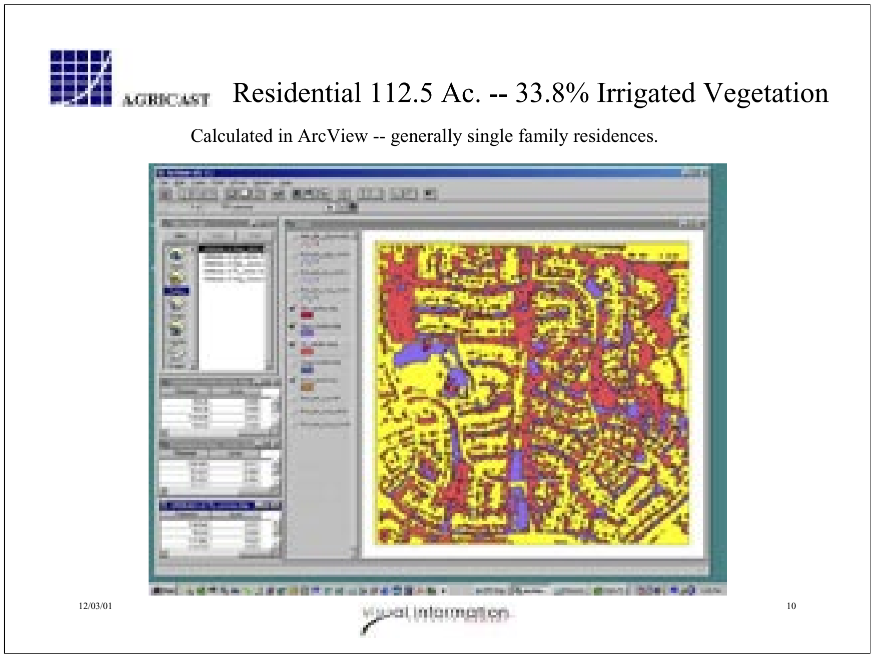


This is one of two ways to calculate areas covered by the various classes resulting from a classification from image processing routines; the other is to convert the raster data to vector and then import the resultant vector file into a Geographic Information System (GIS) such as ArcView (to be discussed later).

When image processing software is used to classify a scene, the routine will produce the statistics relative to each classification. Among the information produced, will be acreage for the entire scene and the area for each of the classified features, generated as a text file. Here the text file has been imported into Excel and edited to show just the area covered by each class.

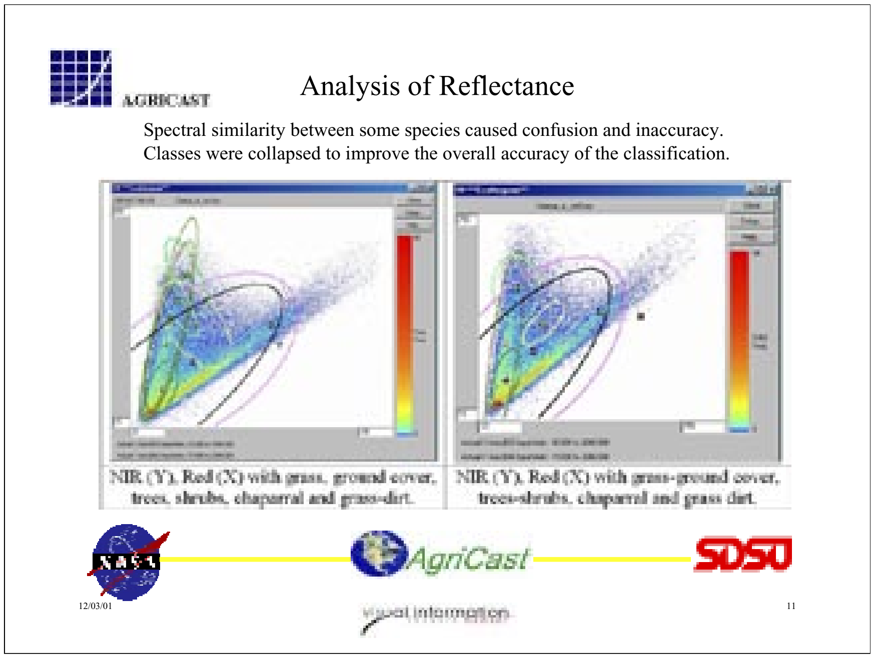
For these areas in Del Mar, California, the single family area was 58 percent irrigated vegetation, largely in large median areas between a row of homes along the west side one street and the next row of homes along the east side of the next street to the west. Or visualized another way, the median area down the center of a city block between two rows of homes.

The condominium area amounted to 43 percent irrigated vegetation. The area was characterized by widely spaced two story condominiums with lush vegetation all around.



Shown here is an example of a more recently developed, very up-scale residential area in Del Mar. This is a composite where polygons representing grass and dirt, trees and shrubs, and grass and ground cover have been imported into ArcView. Yellow represents impervious surfaces such as streets, parking areas and roofs. Blue is grass and ground cover. Red is trees and shrubs. Area was calculated for each set of polygons and added together. This area amounted to 112.5 acres of which 34% was irrigated vegetation.

As to relative accuracy between the two methods, the writer's opinion is statistics generated by the image processing software are more accurate. This is because generation of the vector polygons is sometimes incomplete, particularly when there are interior polygons among larger exterior ones.



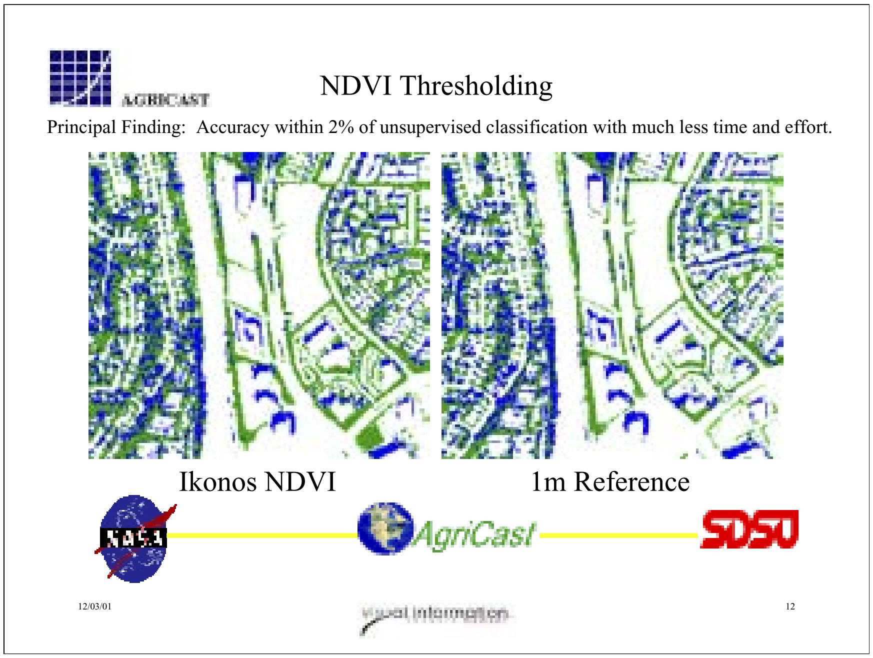
These plots are called scattergrams and are used to evaluate training sites used for a supervised classification. The general rule of thumb is -- keep ellipsis representing the various classes as far apart as possible.

Our first attempt was to classify everything: impervious, chaparral, trees, shrubs, grass and dirt, grass, and then ground cover. The result was misclassification and the reason for the misclassification is shown at the upper left. Look along the vertical axis to see where the ellipsis which represent the spectral reflectance for each class at near infrared overlap one-another.

The solution is shown at the right where these classes were collapsed into more generalized classes of similar reflectance: blue green was used for chaparral, brownish green for tree-shrub, light green for grass and ground cover and yellow for dirt-grass.

This serves to make the point that as good as the sensors are on this new generation, high resolution satellite, there are still limitations. New, hyper-spectral sensors programmed for launch in the two years or so should provide the capability for more robust analysis and better discrimination between plant species.

11



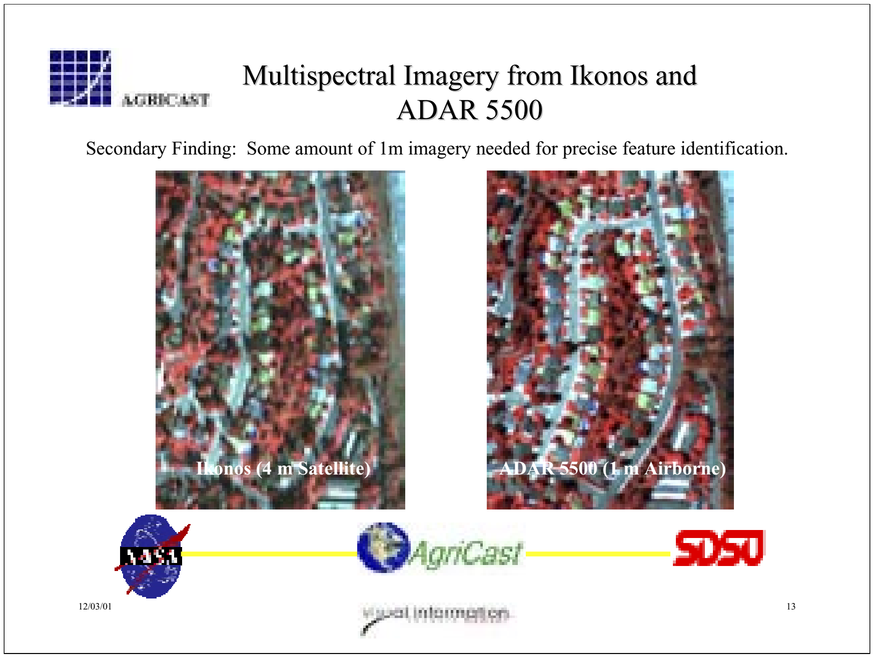
When the limitations of multispectral analysis and classification techniques were noted, we then looked for a less time consuming process to achieve about the same result -- NDVI thresholding.

Normalized Difference Vegetative Index (NDVI) is a classic remote sensing algorithm used to measure healthy vegetation. Essentially this process works by measuring the reflectance of healthy vegetation in the near infrared portion of the infrared band and comparing this to a measurement of dirt or other non-reflective surfaces using the red band.

The formula used in the NDVI algorithm is NIR - Red / NIR + Red. So if we write this as an if / then statement:{ **if** ((NIR-Red/NIR+Red)) >.20 **then** input1 else null } every pixel less than .20 will be null (black) and those above .20 will be shown and accounted for in the statistics. The output is usually gray scale, but other color schemes may be used to show intensity above the threshold value like that above.

The process is quick and simple using Earth Resources Mapping software. Simply enter a cutoff value in the formula, click apply and immediately seethe result. Change the cutoff value up or down, click apply and see that result.

Overall, results from the NDVI Slice technique turned out to be within 2% of the unsupervised classification. We concluded that results from unsupervised classification of a small area may be used to set a threshold which can then be applied over a much larger adjacent area. This will be discussed later.



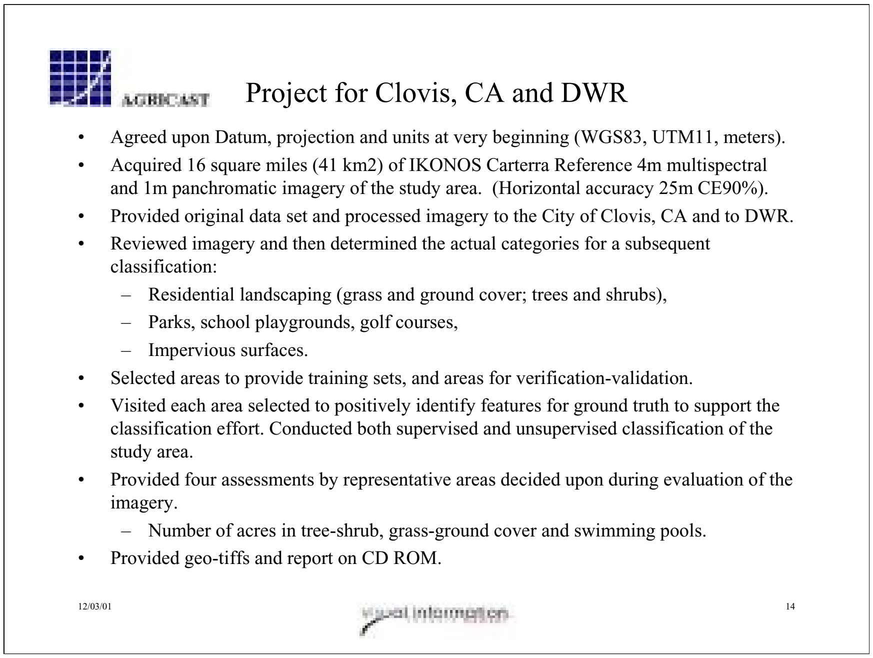
A secondary finding of the work with SDSU was that 1m imagery was needed to precisely identify some features for classification and for verification -validation.

Shown here is a comparison between 4m IKONOS multispectral imagery and the same area as depicted in 1m ADAR aerial imagery. There is no real comparison of cost between one and the other. IKONOS Reference 4m multispectral imagery at 25m horizontal accuracy (90%CE) will cost $29 per km2. ADAR will cost many times that depending upon the total area flown in the subscription.

However, IKONOS offers 1m Panchromatic of the same area for an additional 50% or $43 per km2, so this is the number to be used when comparing to ADAR. The combination is called “pan sharpened” or “1m color”. Examples are shown on the following slide.

Consequently, we will normally propose the full (pan + multi-spectral) data set be used for analysis. If this becomes too costly, then we will look to obtain enough panchromatic to cover the areas identified for detailed analysis, which used to set the threshold for NDVI of the total area. The minimum buy is 100 km2 for either or both IKONOS 4m multispectral and/or 1m panchromatic imagery.

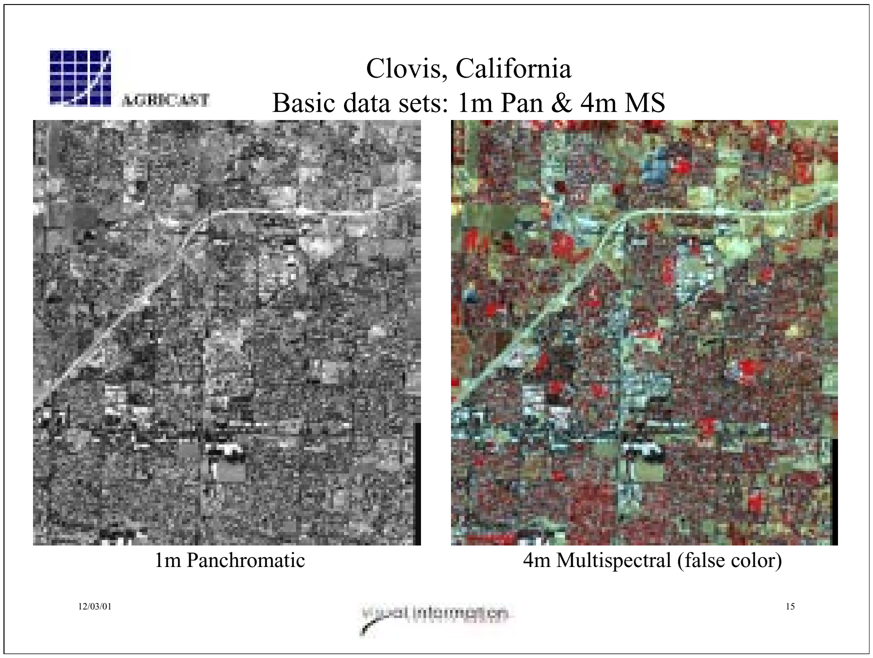
13



The California Department of Water Resources, San Joaquin District and the City of Clovis, California became interested in the project underway at San Diego State University and wished to see how well these techniques and data could be applied to areas in the San Joaquin Valley.

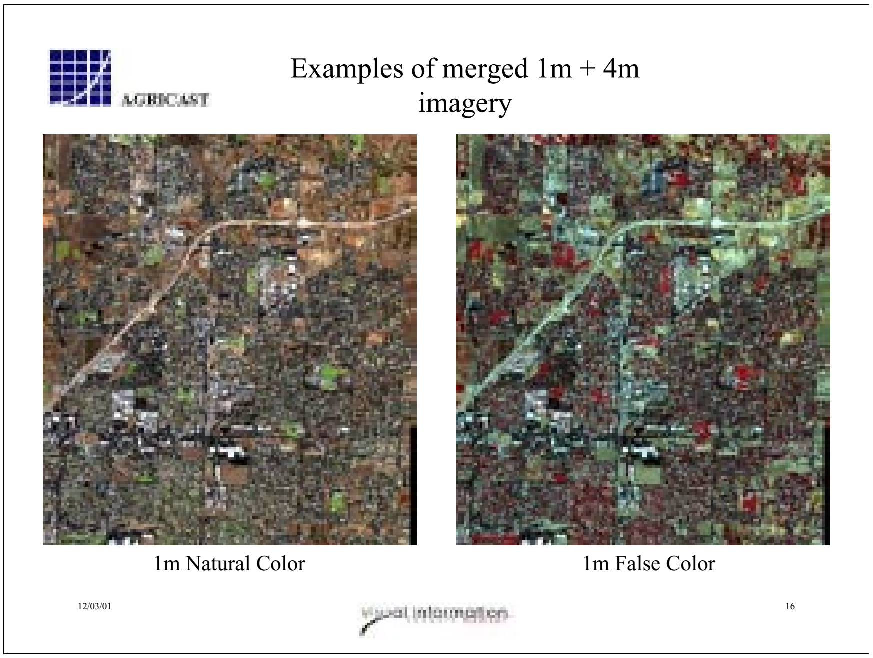
Once on contract, we immediately agreed the areas of interest and on the datum, projection and units of measure being used by Clovis and DWR GIS Departments. But the timing was such that had we ordered imagery at that point, the result would have been a winter scene. Instead it was agreed to place the imagery order for future acquisition in June or July 2001. Actual acquisition occurred on June 20th.

The image data sets were processed and immediately delivered in .tif format for immediate use as base maps. At the same meeting all participants evaluated the scene and selected specific areas for detailed evaluation. The remainder of the project proceeded as summarized on the slide, and described on the following pages.



IKONOS imagery is delivered in 1-meter panchromatic format as shown on the left and/or 4-meter multi-spectral format as shown on the right. The multispectral format consists of four data sets: one for red, green, blue and near infra-red. Shown at upper right is the classic way an image processor will depict an infra-red scene using the near infra-red, red, and green bands. The result is called “false color”.

Healthy vegetation is shown in tones of red with irrigated grass being the brightest. Areas of unirrigated grasslands are shown in green. Urban areas are shown in blue-gray, with some building roofs shown in light gray. The very dark areas are surface water ponds.



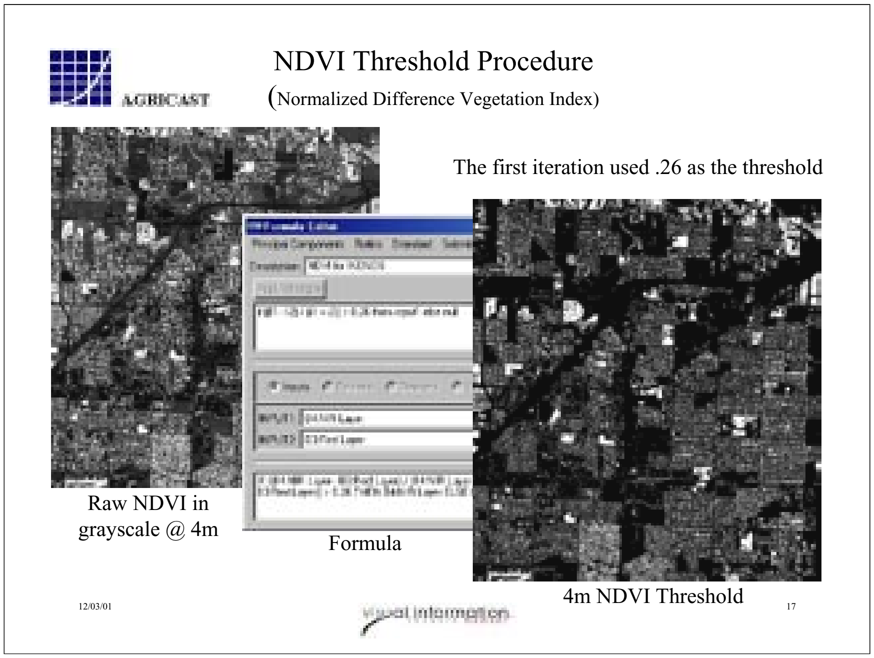
There are other ways to manipulate the bands in multi-spectral imagery. The scene at left is referred to as “natural color” and is accomplished using the red green and blue bands -- not the near infra-red band. At right is the false color scene described on the previous page.

However there is another difference. It is possible and quite easy to merge a panchromatic image of an area with the multi-spectral image of the same area. The result is 1-meter color.

Visually, the depiction is the same as 1m ADAR and is exceptionally useful when small features are to be identified for subsequent classification. But analysis is performed on the 4-meter multi-spectral data set alone and not the panchromatic.

So then it is fair to say that an ADAR derived classification is more accurate because of the discrimination of 1meter pixel versus what is possible using a 4-meter pixel. Work at SDSU showed this difference in accuracy to be about six percent. Then the question becomes one of whether the cost for six points of improved accuracy is worth the added expense?

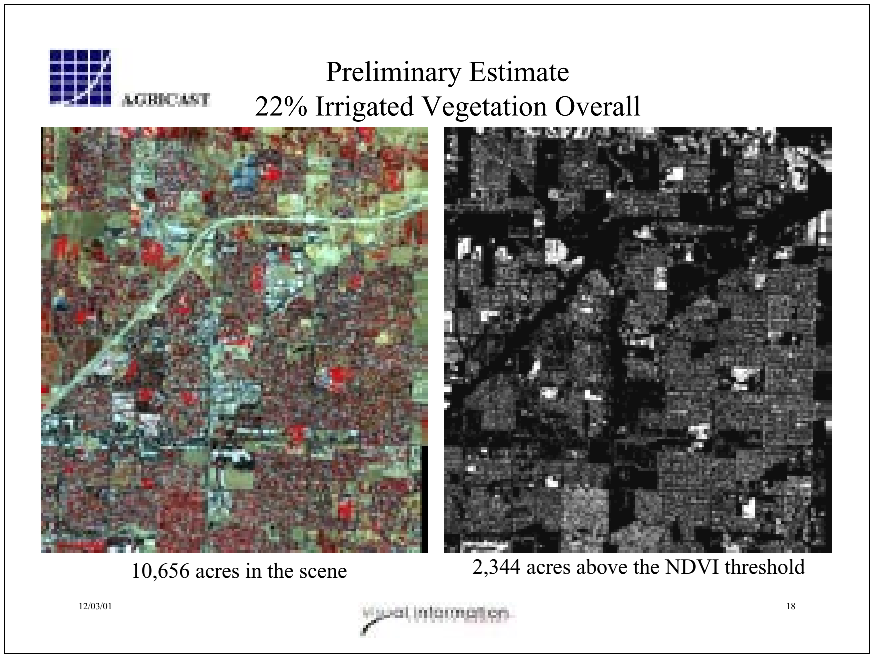
16



Once we agree that accuracy of an satellite imagery derived analysis is not absolute; and that the best that might be expected is about six percent from reference, then the question becomes, is eight percent good enough for the application if this reduces processing cost?

Shown here is an initial application of the NDVI Slice technique to the Clovis area. It is easy to iterate the cut-off value. Simply high-light and change the number in the formula. At that point the "Apply changes" button activates. Click the button and immediately see the result. If it doesn't look right then change the value and try again.

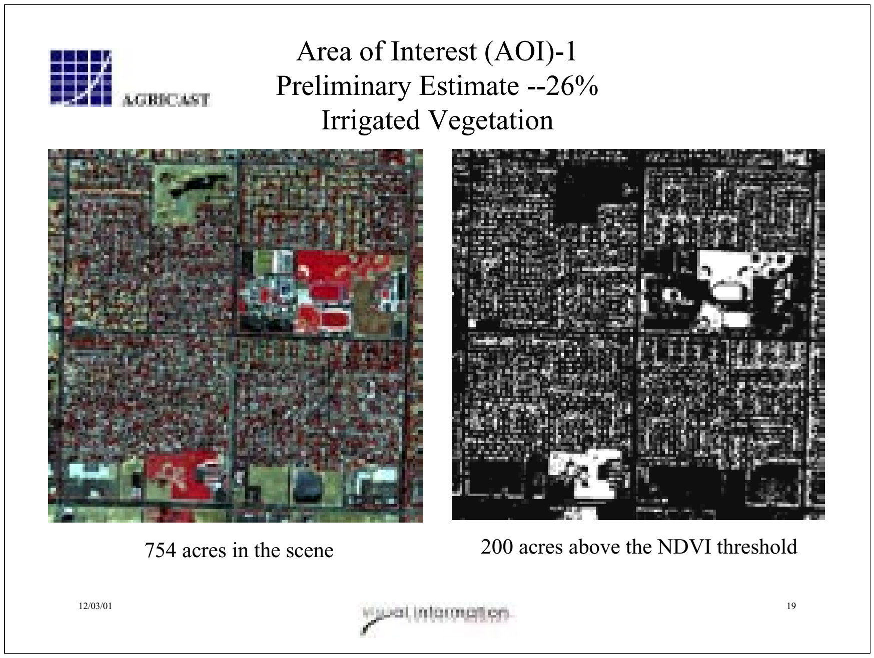
17



Does one value (such as .14 used at Del Mar) fit all types of vegetation and localities. Probably not due to differences in species and climates, even micro climates (or Evapotranspiration Zones).

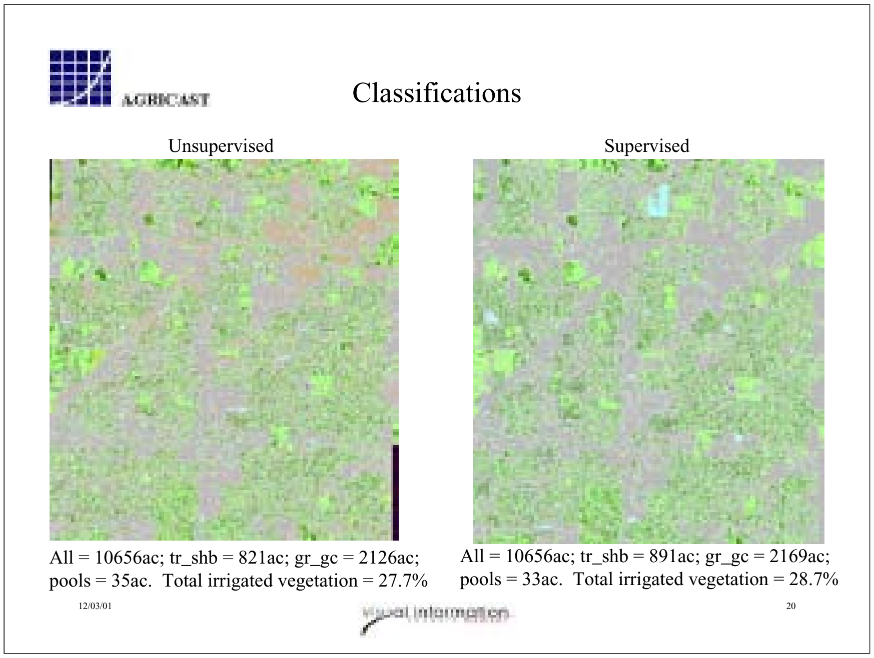
After a number of iterations, .26 was chosen as a test reference for the time being and resulted in 22% of pixels in the scene being above the cut-off value.

18



The next step was to zoom into an area of interest slightly larger than one section or one square mile. Here the preliminary result was about 26% irrigated vegetation using .26 as the cutoff.

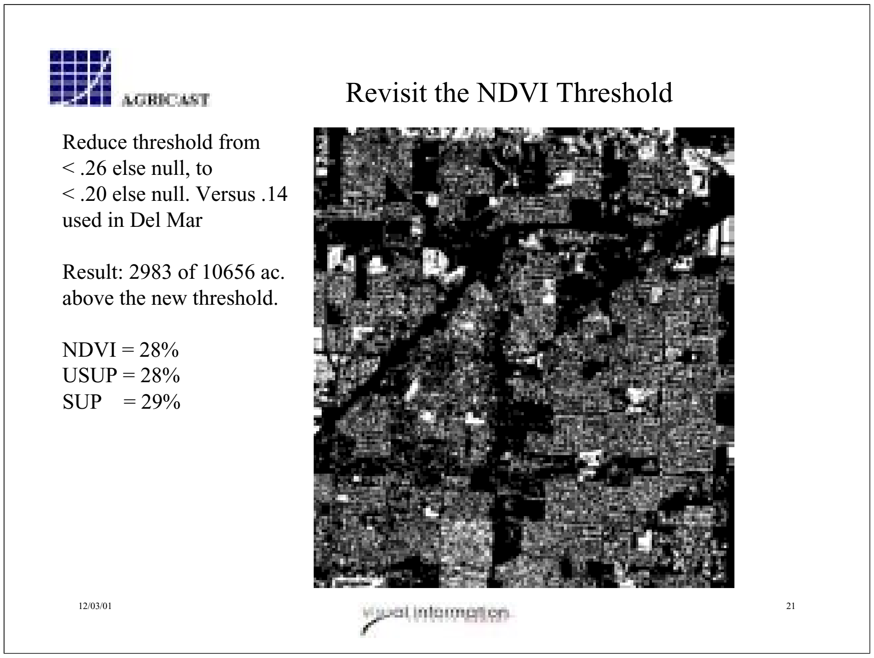
The remainder of the analysis was used to examine other areas in Clovis, CA using unsupervised and supervised classification and then use these results to refine the NDVI cutoff value.



The first step was to perform the unsupervised classification and then use these results as training areas for the supervised classification. The color convention used in both:

gray	impervious (roofs, streets)
lt. brown	bare ground
lt. green	grass & ground cover
dk. green	trees and shrubs
lt. blue	surface water
purple	swimming pools
red	shadow

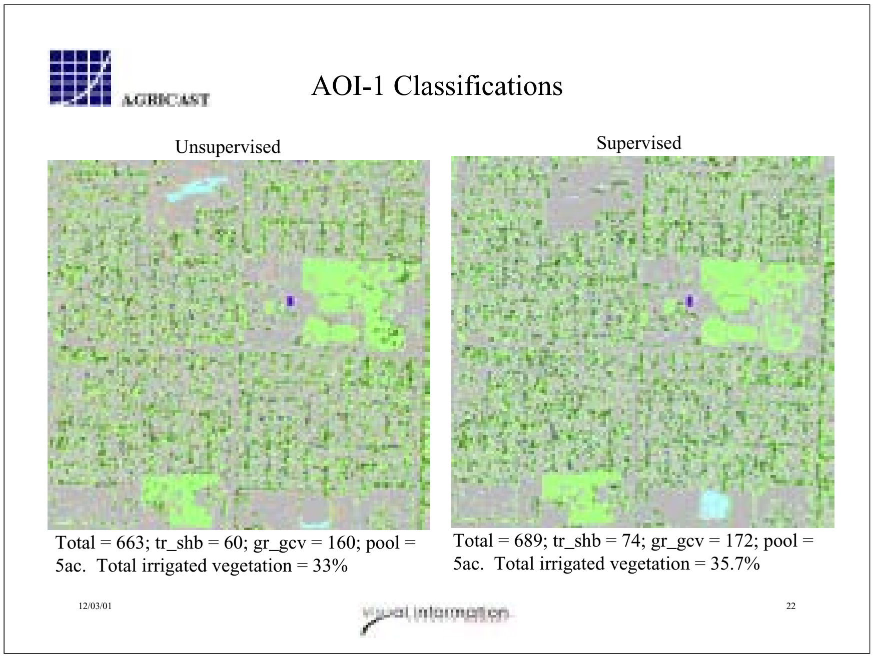
Almost immediately it was noted that there was very little shadow in the scene compared to the Del Mar imagery. This is due to sun elevation at time of year. June acquisition for the Clovis imagery; January acquisition for Del Mar. Additionally the overall results for the unsupervised classification and the supervised classification are almost the same due to the absence of shadow.



With the classified results known and verified, the NDVI threshold was revisited, reduced to .20, and then the statistics were recalculated. This resulted in 2983 pixels in the scene being above cut-off or 28 percent; strong correlation with the classified results.

Sub scenes from the unsupervised, supervised and NDVI classifications were then calculated and compared. These are shown on the following slides.

21



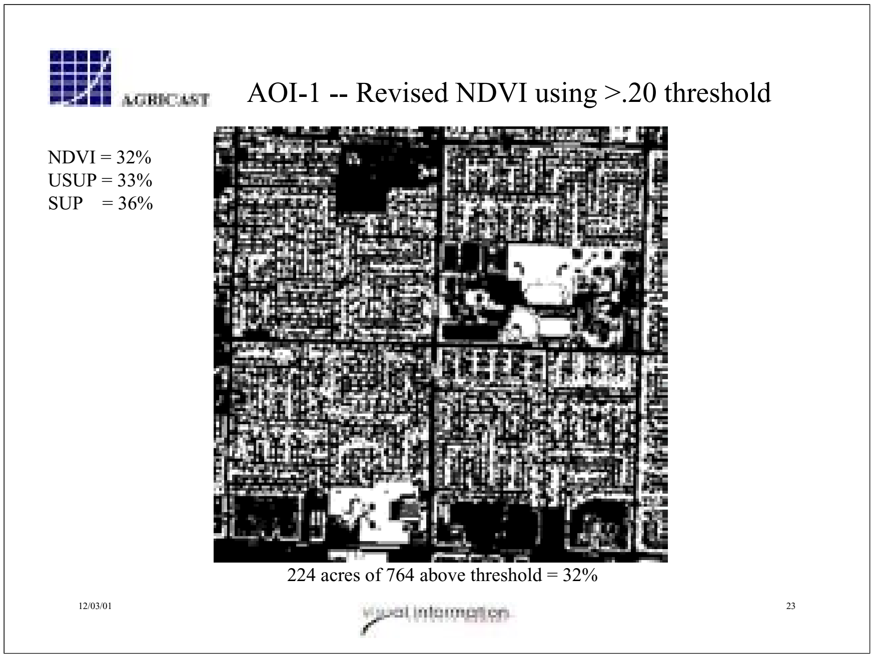
The unsupervised classification is shown at left with the supervised classification shown at right. The first thing that is noticeable between the two is the light blue areas depicting surface water.

The problem with this class was that surface water is as dark a reflectance as some of the asphalt sections of street paving. This was simply a judgement call of which was most important to show -- the paving areas without misclassification or the full extent of the surface water.

Surface water had a very different reflectance compared to swimming pools due to the concrete bottoms which made swimming pools reflect light differently. Pools are shown in purple. The pool at upper right center is an very large pool at a school.

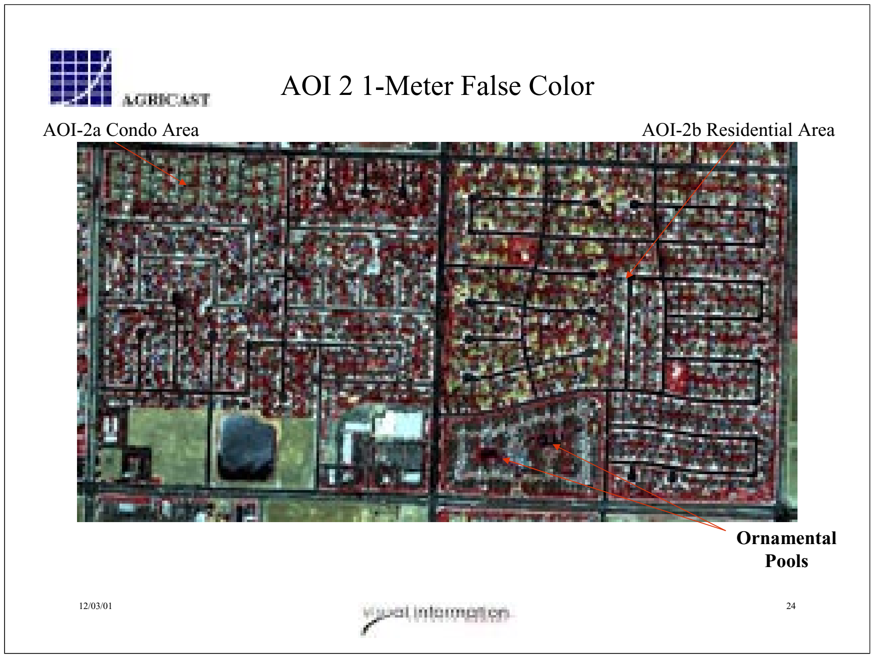
Overall the irrigated vegetation results of the two classifications are very similar:

Unsupervised	33%
Supervised	36%



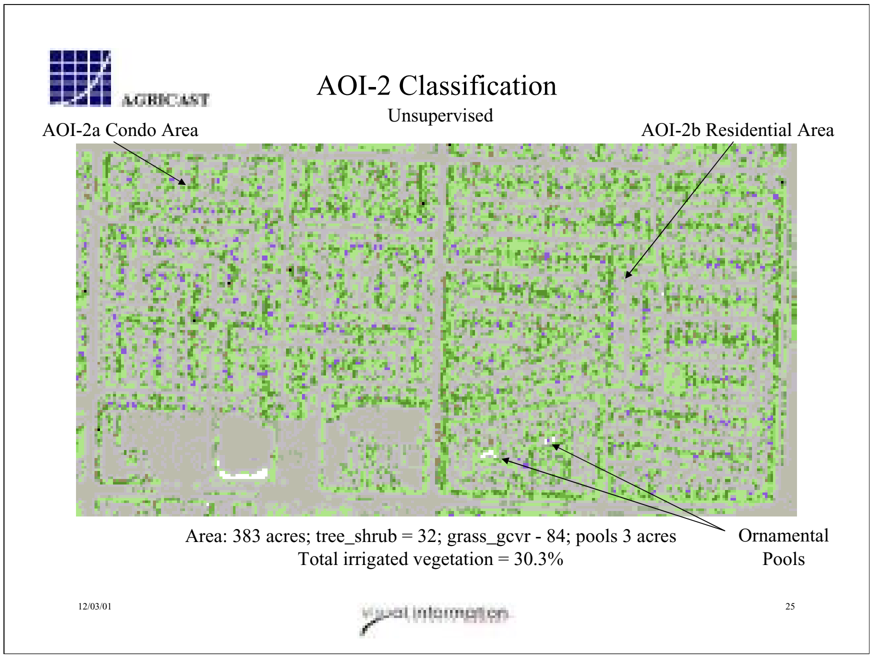
This is the NDVI classification of the same area showing 32% irrigated vegetation. The comparison between this calculation and the classifications is shown on the slide.

23



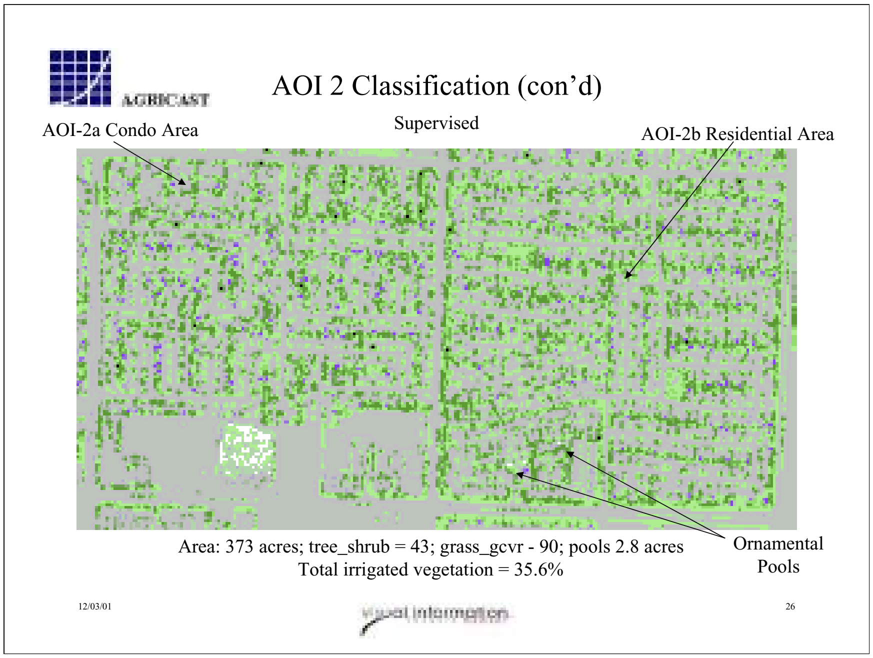
This is a 1-meter false color scene. This area is largely single family residential with the exception of a small condominium area at the upper left. Time of development is older to the left, newer and slightly more up-scale to the right. Two areas of ornamental pools were detected at lower right of center.

24



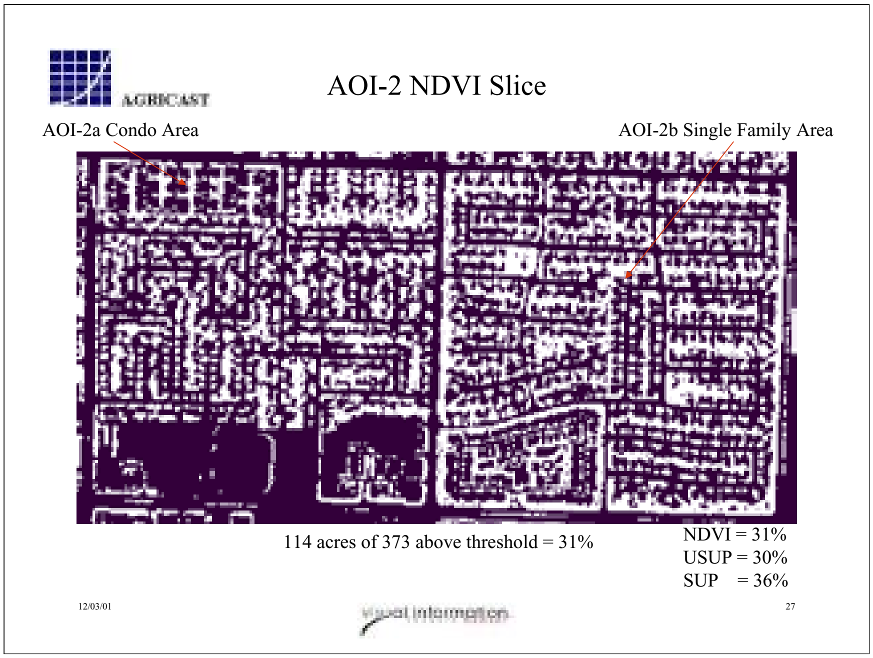
Total irrigated vegetation in this area -- 30.3%. Better classification was achieved for the ornamental ponds here in the unsupervised case than using the supervised routine as shown on the next slide.

25

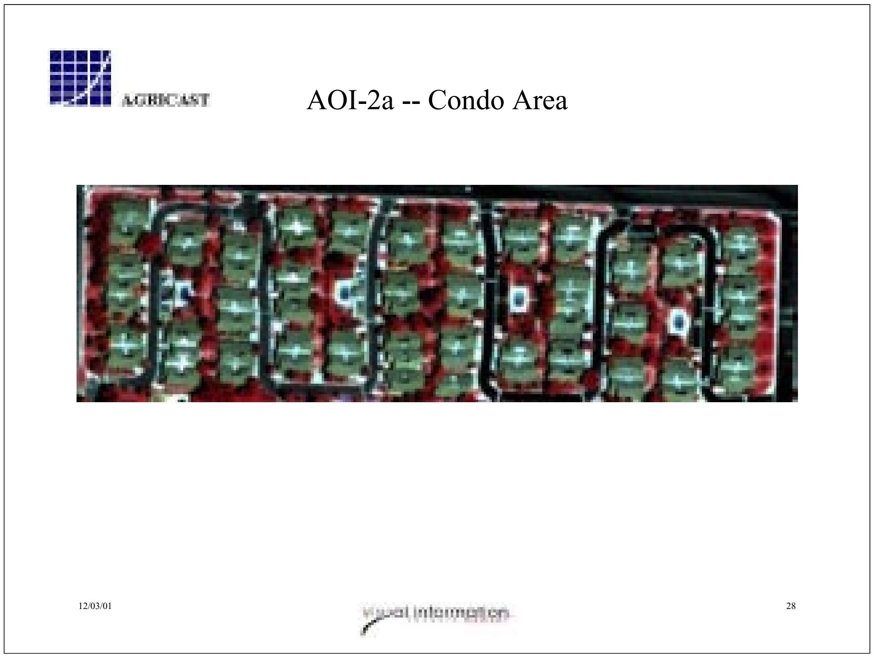


This is the supervised classification of the same area. Total irrigated vegetation using the supervised classification procedures is 35.6%

26

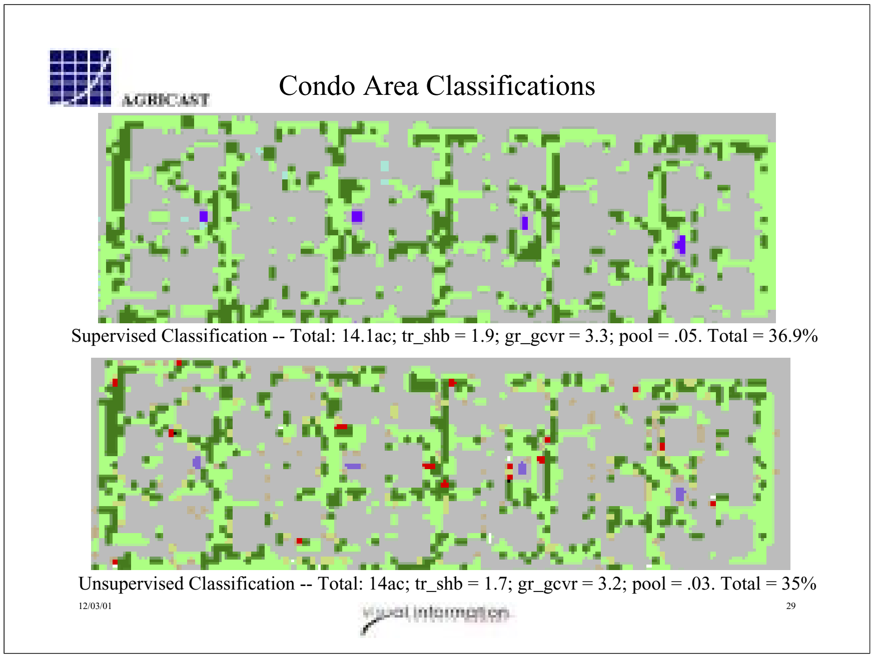


This is the NDVI classification at 31% irrigated vegetation with falls into line with the Unsupervised classification.



This is a zoom into the condominium area shown at the upper left of the previous scene. This is a very small area measuring slightly more than 14 acres. Notice the four swimming pools shown in the scene.

28



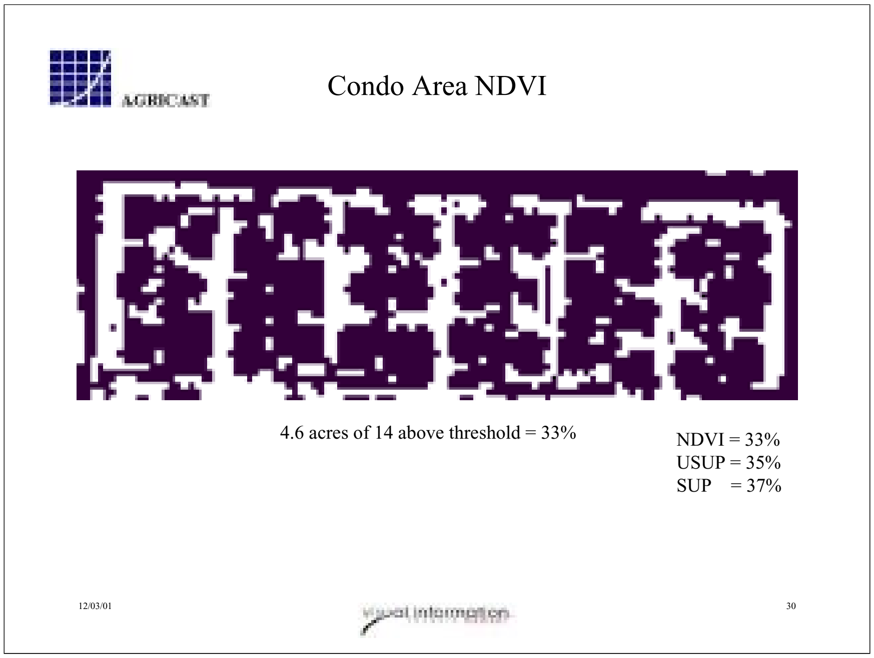
Thus is the supervised classification at top and the unsupervised classification shown at bottom. Calculations between the two are similar. However notice the swimming pools. There are differences in the number of pixels calculated as pools in the two classifications. From left to right:

Pool 1	Supervised 2	Unsupervised 2
Pool 2	Supervised 4	Unsupervised 2
Pool 3	Supervised 2	Unsupervised 2
Pool 4	Supervised 3	Unsupervised 2

Eleven pixels in the supervised classification compared to eight pixels in the unsupervised classifications accounts for the differences in the calculation for swimming pool area.

We were also able to account for some shadow in the unsupervised classification as shown in red -- too small an amount to attempt in the supervised classification for fear of misclassification of the adjacent trees.

29



This is the NDVI classification for the Condominium area with the comparison between this method and the classifications shown on the slide. The results from the NDVI classification compare favorably with the unsupervised classification.

AGRICAST

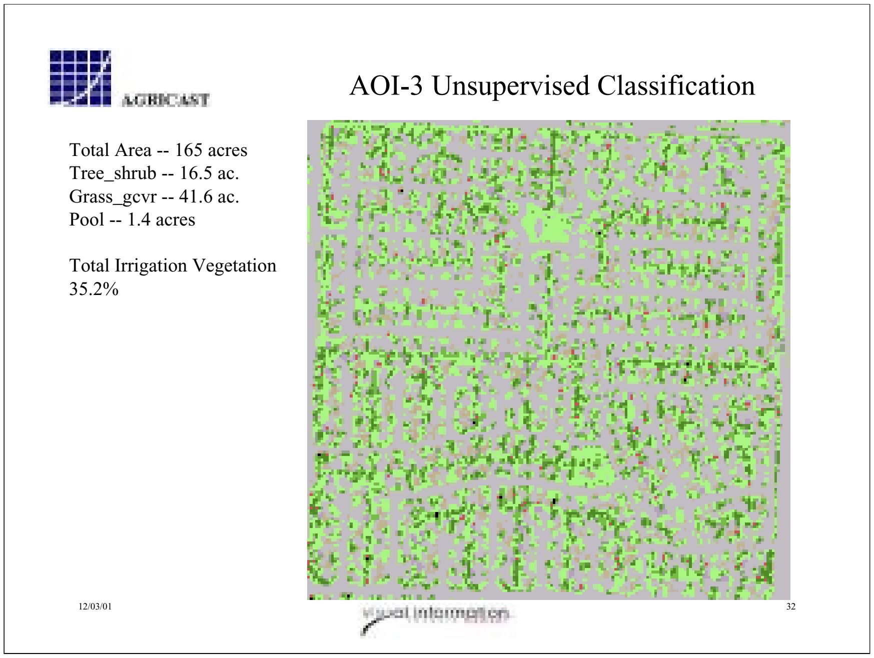
AOI-3, Single Family

12/03/01

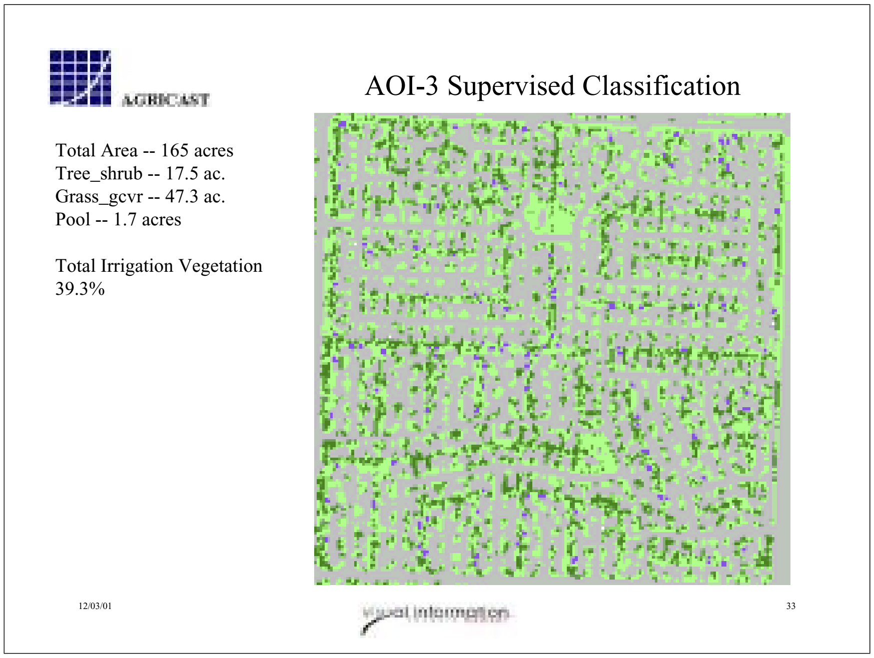
visual information

31

This is a single-family residential area slightly up-scale from those previously shown and very similar to the newly developed upscale area in Del Mar where the percentage of irrigated vegetation was 34 percent.

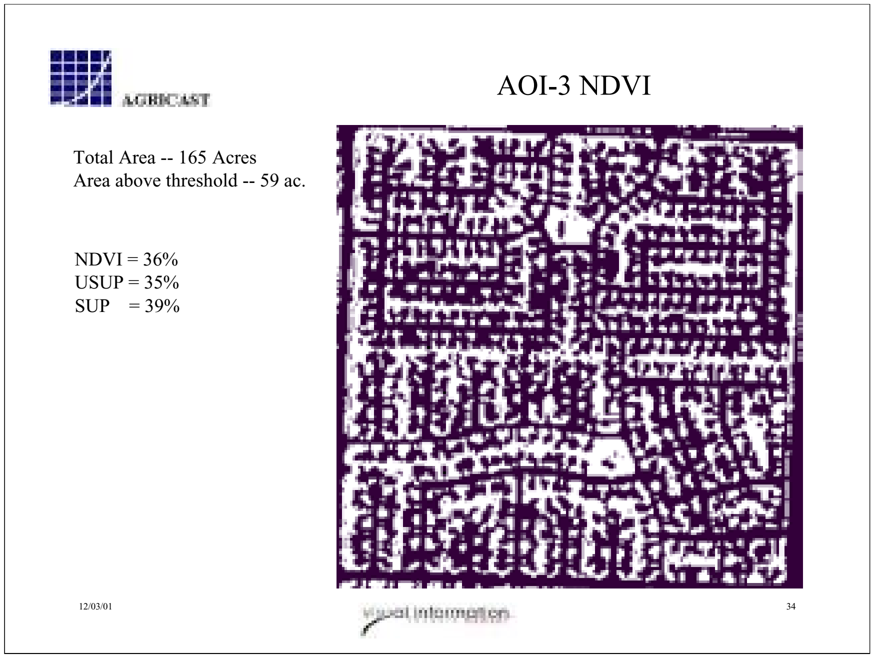


This is the unsupervised classification showing about 1.4 acres in swimming pools and an overall 35 percent of irrigated vegetation.



Compared to the supervised classification showing 1.7 acres in swimming pools and an overall 39 percent in irrigated vegetation.

33



The NDVI classification falls between the previous classifications but, as in past cases, and as in Del Mar, tends to favor the unsupervised classification. The comparison between methods is shown on the slide.

34